Course : Data Science, the fundamentals

Synthesis course - 2d - 14h00 - Ref. DTX
Price : 2020 CHF E.T.

4,3 / 5

BEST

A major strategic challenge for organizations, data science uses mathematical tools to reveal the behavior of data and analyze the events they describe. This course covers the fundamentals of data science and provides an introduction to data analysis.

Teaching objectives

At the end of the training, the participant will be able to:

- Understand the basic principles of data science and how to organize the process
- Understand the application of data science to solve problems and its limitations
- Develop your ability to analyze and interpret figures through graphical representation
- Understand how to use data science tools and develop models for professional use
- Opening up to the challenge of exploiting data in a competitive and continuous improvement context
- Understanding the organization and infrastructure for data science services and projects

Intended audience

IS managers, data analysis project managers, statistical research managers.

Prerequisites

No special knowledge required.

PARTICIPANTS

IS managers, data analysis project managers, statistical research managers.

PREREQUISITES

No special knowledge required.

TRAINER QUALIFICATIONS

The experts leading the training are specialists in the covered subjects. They have been approved by our instructional teams for both their professional knowledge and their teaching ability, for each course they teach. They have at least five to ten years of experience in their field and hold (or have held) decision-making positions in companies.

ASSESSMENT TERMS

The trainer evaluates each participant's academic progress throughout the training using multiple choice, scenarios, hands-on work and more.
Participants also complete a placement test before and after the course to measure the skills they've developed.

TEACHING AIDS AND TECHNICAL RESOURCES

• The main teaching aids and instructional methods used in the training are audiovisual aids, documentation and course material, hands-on application exercises and corrected exercises for practical training courses, case studies and coverage of real cases for training seminars.
• At the end of each course or seminar, ORSYS provides participants with a course evaluation questionnaire that is analysed by our instructional teams.
• A check-in sheet for each half-day of attendance is provided at the end of the training, along with a course completion certificate if the trainee attended the entire session.

Practical details

Hands-on work

Guided practice of the fundamentals through exercises. MCQs and summary tables help you to position yourself.

Teaching methods

During this summary course, the trainer gives demonstrations that each participant reproduces to put the main concepts into practice.

TERMS AND DEADLINES

Registration must be completed 24 hours before the start of the training.

ACCESSIBILITY FOR PEOPLE WITH DISABILITIES

Do you need special accessibility accommodations? Contact Mrs. Fosse, Disability Manager, at psh-accueil@orsys.fr to review your request and its feasibility.

Course schedule

1 What is data science?

- Fundamentals: big data, data lake, data mining, artificial intelligence, machine and deep learning, text mining.
- New challenges: the emergence and multiplication of new data sources.
- Take into account data heterogeneity, real-time flows and the explosion of data volumes.
- The big data technological ecosystem.
- Demystify the world of data science: descriptive, predictive and prescriptive analysis.
- The job, tools and methods of the data scientist.
- Introduction to machine learning, supervised analysis and unsupervised analysis.
- Notions of over- and underlearning.

Demonstration

Use cases for data science in a business value chain (customer behavior, product offering, etc.).

2 Data science methods and models

- Data collection, preparation and exploration.
- The importance of the data quality approach (cleaning, transforming, enriching).
- Definition of metrics.
- Basic statistical methods.
- The main classes of supervised algorithms: decision trees, K-nearest neighbors, regression, Naive Bayes.
- The main classes of unsupervised algorithms: clustering, PCA, CAH, neural networks.
- Text mining and other families of algorithms.

Storyboarding workshops

Simple analyses with R or Python to illustrate supervised (regression and classification) and unsupervised (clustering, segmentation and anomaly detection) analysis techniques.

3 Graphical representation and data retrieval

- R and Python statistical analysis languages.
- Their development environments (R-Studio, Anaconda, PyCharm) and libraries (Pandas, machine learning).
- DataViz tools (Power BI, Qlik, tableau, etc.).
- Data modeling: representation of processes, flows, controls and conditions.
- Data modeling: tools (Orange, Power BI).
- Communicate results through data storytelling: organize visuals (diagrams, rankings, maps).
- Communicating results through data storytelling: conveying the meaning of results.

Storyboarding workshops

Exercises in graphical data exploration, analyzing the position and extent of data (clouds, histograms, etc.).

4 Modeling a data science problem

- Summary of the process.
- Analysis of two business cases, customer relations and fraud detection for example, but others are also possible.
- Case study 1: customer relations in the insurance industry.
- Cibler les campagnes marketing. Comprendre les causes d'attrition client. Quels produits pour quels clients ?
- Case study 2: fraud detection.
- Compare research using classical statistics and data mining.
- Detection by supervised method. Unsupervised detection.

Case study

Practical application of the storytelling method to business cases.

Dates and locations

REMOTE CLASS
2026 : 24 Mar., 26 May, 23 June, 29 Sep., 13 Oct., 8 Dec.